Int. J. Contemp. Math. Sci., Vol. 2, 2007, no. 12, 571 - 580

Finding the Piecewise Linear Frontier Production Function in DEA with Interval Data

G. R. Jahanshahloo and F. Hosseinzadeh Lotfi[1]

Department of Mathematics, Science and Research Branch
Islamic Azad University, Tehran, Iran

F. Rezai Balf

Department of Mathematics
Islamic Azad University, Qaemshahr, Iran

P. Zamani

Department of Mathematics
Islamic Azad University, Karaj, Iran

Abstract

Data envelopment analysis (DEA) is a mathematical programming technique for identifying efficiency scores of decision making units (DMUs). Since DEA models cannot present efficient frontiers of PPS, in order to do this, we introduce a method for identifying efficient frontier for DMUs with interval data.

Keywords: Data envelopment analysis, Interval data, Efficiency, Frontier

1. Introduction

Data envelopment analysis (DEA) is a nonparametric method for evaluating efficiency of systems with multiple inputs and multiple outputs. Its goal is to classify the decision making unit (DMU) in to two classes: efficient or inefficient ones. Based on inputs and outputs of the units, DEA forms efficient surfaces. If a DMU lies on the surface, it is efficient; otherwise, it is inefficient. However, uncertainty such as a measurement error should be incorporated in observed data. The original DEA models assume that inputs and outputs are measured by exact values on a ratio scale.

---

[1] Corresponding author: Farhad Hosseinzadeh Lotfi, E-mail: hosseinzadeh_lotfi@yahoo.com, Tel: 98-21-4484172, Fax: 98-21-44804172

Recently, Hosseinzadeh lotfi et al. [8] addressed the problem of imprecise data in DEA in its general form. Imprecise data means that some data are known only to the extent that the true values lie within prescribed bounds while other data are known only to satisfy certain relations.

Korhonen [11] tried to provide the Decision Maker (DM) an interactive method which allows him/her to incorporate performance information in to the efficient frontier analysis by enabling him to make a free search on efficient frontier. Indeed, he wants to provide the DM all references of an inefficient DMU, enabling him/her to choose the most preferable unit as reference. But finding all references of an inefficient DMU is not an easy job. Also, efficient surfaces are useful in analyzing DEA efficient DMUs and nondominated solutions (Pareto solutions) in multiobjective programming.

Not many papers have been written on the subject of "finding efficient frontier". Jahanshahloo et al. [10] proposed a method to obtain efficient frontier by using 0-1 integer programming. Although, solving 0-1 integer programming problem needs more computational efforts, but for large n (number of DMUs), solving a little amount of integer problems may be more preferable than solving n linear programs. This is especially important in situations where the number of efficient facets is very less than n e.g. when $m + s$ (number of inputs and outputs) is very less than n.

In this paper, we try to develop the above mentioned way to obtain the efficient frontier for DMUs with interval data.

The current article proceeds as follows: In section 2, we review DEA models for dealing with interval data. In section 3, we obtain the efficient frontier for DMUs with interval data. In section 4, a numerical example is considered. A concluding section summarizes our main results.

2. Preliminaries

Production possibility set (PPS) is defined as the set of all inputs and outputs of a system in which inputs can produce outputs. For evaluating the efficiency of DMU_0 $(0 \in \{1, \ldots, n\})$, the input-oriented model (2.1) is applied. Model (2.2) is the output-oriented model.

$$\begin{array}{ll} \min & \theta \\ s.t & X\lambda \leq \theta x_0, \\ & Y\lambda \geq y_0, \\ & \lambda \in \Lambda, \end{array} \tag{2.1}$$

and

$$\begin{array}{ll} \max & \phi \\ s.t & X\lambda \leq x_0, \\ & Y\lambda \geq \phi y_0, \\ & \lambda \in \Lambda, \end{array} \tag{2.2}$$

where $X = (x_j)$ and $Y = (y_j)$, $j = 1, \dots, n$. The above mentioned models are called CCR if $A = \{\lambda | \lambda \geq 0\}$ and BCC if $\Lambda = \{\lambda | \lambda \geq 0, e\lambda = 1\}$. The PPS of CCR model is T_c and the PPS of BCC model is T_v. The models (2.1) and (2.2) are called envelopment form. T_c and T_v are depicted in Fig 1.

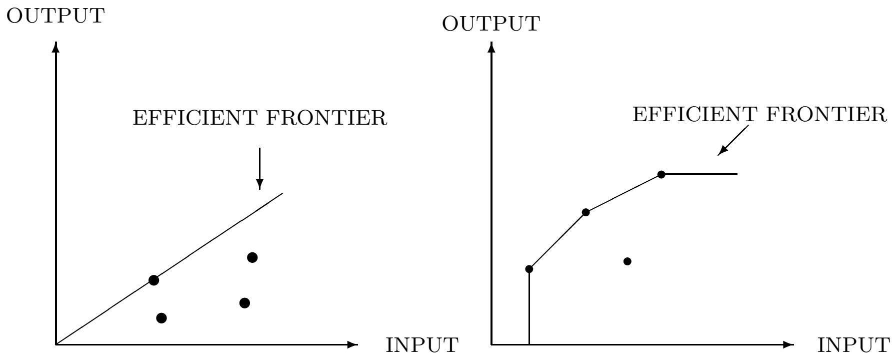

Fig 1. Illustration of T_c and T_v

Suppose we have n DMUs, where each DMU$_j$ ($j \in J, J = \{1, \dots, n\}$) consumes m inputs $x_j^t = (x_{1j}, \dots, x_{mj})$ to produce s outputs $y_j^t = (y_{1j}, \dots, y_{sj})$. We assume $y_j \geq 0$, $x_j \geq 0$, $y_j \neq 0$, $x_j \neq 0$. Unlike the original DEA model, we assume further that the levels of inputs and outputs are not known exactly, the true input and output data known to lie within bounded intervals, i.e. $x_{ij} \in [x_{ij}^L, x_{ij}^U]$ and $y_{rj} \in [y_{rj}^L, y_{rj}^U]$ with upper and lower bounds of the intervals given as constants and assumed strictly positive. Consider to the following graph:

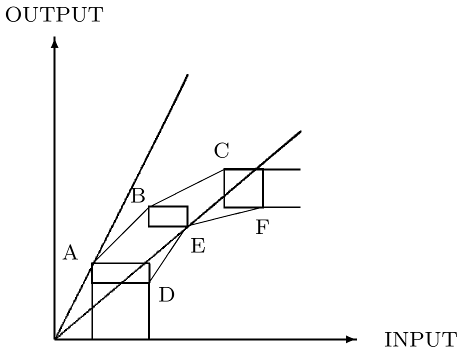

Fig 2. Interval Data in T_c and T_v

in which the points A, B and C have the least input and most output, and the points D, E and F have the most input and least output. T_v and T_c have been specified for both groups (Fig 2).

Now, consider to the following model:

$$\begin{array}{lll} \min & \theta & \\ s.t & \sum_{j=1}^{n} \lambda_j [x_{ij}^L, x_{ij}^U] \leq \theta [x_{i0}^L, x_{i0}^U], & i = 1, \ldots, m \\ & \sum_{j=1}^{n} \lambda_j [y_{rj}^L, y_{rj}^U] \geq [y_{r0}^L, y_{r0}^U], & r = 1, \ldots, s, \\ & \lambda \in \Lambda. & \end{array} \tag{2.3}$$

Which is a CCR/BCC model with the input-oriented associated interval data. Now consider to the following linear programming problem:

$$\begin{array}{lll} \max & z = \sum_{j=1}^{n} c_j x_j & \\ s.t & \sum_{j=1}^{n} a_{ij} x_i \leq b_i, & i = 1, \ldots, m \\ & x_j \geq 0, & j = 1, \ldots, n, \end{array} \tag{2.4}$$

with $c_j \in [c_j^L, c_j^U]$, $a_{ij} \in [a_{ij}^L, a_{ij}^U]$ and $b_i \in [b_i^L, b_i^U]$.

One way to obtain the solutions of the above model is to solve the following problem:

$$\begin{array}{lll} \max & z = \sum_{j=1}^{n} c_j x_j & \\ s.t & \sum_{j=1}^{n} a_{ij} x_i \leq b_i, & i = 1, \ldots, m \\ & c_j^L \leq c_j \leq c_j^U, & j = 1, \ldots, n \\ & a_{ij}^L \leq a_{ij} \leq a_{ij}^U, & i = 1, \ldots, m, j = 1, \ldots, n \\ & b_i^L \leq b_i \leq b_i^U, & i = 1, \ldots, m \\ & x_j \geq 0, & j = 1, \ldots, n. \end{array} \tag{2.5}$$

The difficulty of mentioned method is that in the problem (2.5) a_{ij}, b_i and $c_j (i = 1, \ldots, m), (j = 1, \ldots, n)$ are variable. Another method to solve (2.4) with the mentioned conditions in it is to choose quantities of a_{ij}, b_i and c_j which the objective function to maximize or minimize and then prove that the optimal objective value changes in interval $[z^L, z^U]$.

$$\begin{array}{rl} Z^L = \max & C^L x \\ s.t & A^U x \leq b^L \\ & x \geq 0 \end{array} \tag{2.6}$$

$$\begin{array}{rl} Z^U = \max & C^U x \\ s.t & A^L x \leq b^U \\ & x \geq 0 \end{array} \tag{2.7}$$

Theorem 1 *The optimal objective value in problem (2.4) with the conditions in it changes in interval* $[Z^L, Z^U]$.

Proof. First we prove that $Z^L \leq Z^*$, proof of another inequality is similarly. For this, first we prove that the objective function in problem (2.6) is less or equal than (2.4) and the feasible region of problem (2.6) is the subset of problem (2.4). According to $c_j^L \leq c_j$, it is obvious that $C^L x \leq Cx$. But to prove the feasible region, since $a_{ij} \leq a_{ij}^U$ and $b^L \leq b$, we have:

$$\begin{aligned} &\forall x \qquad A^U x \leq b^L \\ &\textstyle\sum_{j=1}^n a_{ij}x_j \leq \sum_{j=1}^n a_{ij}^U x_j \leq b^L \leq b \quad \Rightarrow \quad Ax \leq b. \quad \blacksquare \end{aligned}$$

Now we return to the problem (2.3). In this case, the efficiency can be interval. The upper limit of interval efficiency is obtained from the optimistic viewpoint and the lower limit is obtained from the pessimistic viewpoint.

Suppose that we want to find the maximum rate of efficiency of DMU$_0$, so the evaluated unit should have the least input and most output and other unit's vice-versa. Now we have:

$$\begin{aligned} \theta^U \ &= \min \quad \theta \\ s.t \quad & \textstyle\sum_{\substack{j=1 \\ j\neq 0}}^n \lambda_j x_{ij}^U + \lambda_0 x_{i0}^L \leq \theta x_{i0}^L, \qquad i = 1, \dots, m, \\ & \textstyle\sum_{\substack{j=1 \\ j\neq 0}}^n \lambda_j y_{rj}^L + \lambda_0 y_{r0}^U \geq y_{r0}^U, \qquad r = 1, \dots, s, \\ & \lambda \in \Lambda. \end{aligned} \tag{2.8}$$

Note that we can not use the previous method of obtaining maximum or minimum objective function value and its reason is that in previous case, the right hand side vector (b_i) was independent from the left values (a_{ij}). But in this case x_{io} is one of the quantities of x_{ij}. Now imagine that we are seeking the minimum rate of efficiency of DMU$_0$. Similarly, the unit under evaluation should have the most input and least output and other unit's vice-versa. i.e:

$$\begin{aligned} \theta^L = \min \quad & \theta \\ s.t \quad & \textstyle\sum_{\substack{j=1 \\ j\neq 0}}^n \lambda_j x_{ij}^L + \lambda_0 x_{i0}^U \leq \theta x_{i0}^U, \qquad i = 1, \dots, m \\ & \textstyle\sum_{\substack{j=1 \\ j\neq 0}}^n \lambda_j y_{rj}^U + \lambda_0 y_{r0}^L \geq y_{r0}^L, \qquad r = 1, \dots, s \\ & \lambda \in \Lambda \end{aligned} \tag{2.9}$$

The dual of above models, multiplier model, is as follows:

$$\begin{aligned} \theta^U = \max \quad & u^t y_0 \\ s.t \quad & u^t y_j^L - v^t x_j^U \leq 0, \qquad j = 1, \ldots, n, j \neq 0 \\ & u^t y_0^U - v^t x_0^L \leq 0 \\ & v^t x_0^L = 1 \\ & u \geq 0, v \geq 0 \end{aligned} \tag{2.10}$$

$$\begin{aligned} \theta^L = \max \quad & u^t y_0 \\ s.t \quad & u^t y_j^U - v^t x_j^L \leq 0, \qquad j = 1, \ldots, n, j \neq 0 \\ & u^t y_0^L - v^t x_0^U \leq 0 \\ & v^t x_0^U = 1 \\ & u \geq 0, v \geq 0 \end{aligned} \tag{2.11}$$

Theorem 2 *The optimal objective of (2.3) will change in interval* $[\theta^L, \theta^U]$ *[6].*

3. Identifying the equations of efficient frontier

In this section, we propose a method to produce efficient frontier of T_v. Producing the efficient frontier of T_c is alike. Since the optimal solution of the BCC multiplier model gives the coefficient of a supporting hyperplane of T_v, in each iteration of proposed algorithm by manipulation of the BCC multiplier model, we identify the equation of a supporting hyperplane of T_v at maximum number of DMUs, which gives us an efficient facet of T_v. First, we take notice of the DMUs with the least input and most output and then the DMUs with the most input and least output are considered. Now, we note to the former case:

Let F_k be the equation of supporting hyperplane that is produced in k-th iteration of algorithm, an G_k be defined as follows:

$$G_k = \{\text{DMU}_j \mid \text{DMU}_j \text{ Lies on } F_k\}.$$

Then an algorithm for identifying efficient frontier of T_v for this case is as follows:

Step 0. Let $k = 1$.

Step 1. Solve the following problem:

$$
\begin{array}{lll}
(P_k)\ \max & \sum_{j=1}^{n} P_j & (3.1) \\
s.t & \sum_{r=1}^{s} u_r y_{rj}^{U} - \sum_{i=1}^{m} v_i x_{ij}^{L} + u_0 \leq 0, & j = 1, \ldots, n \\
 & \sum_{r=1}^{s} u_r y_{rj}^{U} - \sum_{i=1}^{m} v_i x_{ij}^{L} + u_0 \geq -M(1 - P_j), & j = 1, \ldots, n \\
 & \sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i \geq 1 & \\
 & \sum_{i \notin G_j} P_i \geq 1, & j = 1, \ldots, k-1 \\
 & P_j \in \{0, 1\}, & j = 1, \ldots, n \\
 & u, v \geq 0 &
\end{array}
$$

Where M is a large positive number and 0-1 variable P_j equals 1, if DMU_j lies on the optimal supporting hyperplane identified by the model, $(\sum_{r=1}^{s} u_r y^U - \sum_{i=1}^{m} v_i x^L + u_0 = 0)$, and 0 otherwise. In this case when $P_0 = 0$ constraint (2) for $j = 0$ is superfluous.

Optimal solution of above model is equation of a supporting hyperplane of T_v, F_k, which identify an efficient facet of it, this hyperplane is different form those hyperplanes that introduced in previous iterations due to the constraint (4). Hence, it is the equation of a new efficient facet of T_v.

Let $G_k = \{\ \mathrm{DMU}_j \mid \mathrm{DMU}_j \text{ Lies on } F_k\}$, $k = k + 1$ and iterate step 1. But, if the above model is infeasible, all of the strong efficient facets (efficient frontier) of T_v, have been produced and algorithm is terminated.

Theorem 3 *When the model P_k is infeasible all of the strong efficient facets of T_v have been produced, so the algorithm is terminated.*

The above algorithm for the next case is exactly like the first, but there is a little difference, which the P_k will be as follows:

$$
\begin{array}{lll}
(P_k')\ \max & \sum_{j=1}^{n} P_j & (3.2) \\
s.t & \sum_{r=1}^{s} u_r y_{rj}^{L} - \sum_{i=1}^{m} v_i x_{ij}^{U} + u_0 \leq 0, & j = 1, \ldots, n \\
 & \sum_{r=1}^{s} u_r y_{rj}^{L} - \sum_{i=1}^{m} v_i x_{ij}^{U} + u_0 \geq -M(1 - P_j), & j = 1, \ldots, n \\
 & \sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i \geq 1 & \\
 & \sum_{i \notin G_j} P_i \geq 1, & j = 1, \ldots, k-1 \\
 & P_j \in \{0, 1\}, & j = 1, \ldots, n \\
 & u, v \geq 0 &
\end{array}
$$

4. Numerical example

Consider the following DMUs, where x_j^L and x_j^U show lower and upper bound for inputs, respectively. A similar notation is used for outputs (Table 1).

Table 1. The lower and upper bound for inputs and outputs

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
x_j^L	15.51	5.99	8.64	9.66	10.76	8.50	13.89	11.47	9.84	7.00	10.19	10.45	9.22	11.63	9.20	6.17	11.22	12.77
x_j^U	16.54	19.79	23.64	17.37	20.85	19.51	16.29	19.70	17.41	18 .73	14.22	13.56	18.05	14.68	17.25	25.10	19.91	12.77
y_j^L	6.10	8.58	8.20	10.23	9.34	2.97	8.95	10.03	7.34	4.59	6.88	9.82	9.30	10.22	8.19	6.95	8.25	8.33
y_j^U	22.57	19.49	19.47	15.14	19.45	23.60	20 .04	18.26	19.67	24.99	19.77	15.99	17.99	19.15	18.48	21.66	19.72	17.04

First, we obtain the equations of efficient facet for DMUs with least input and most output. Using the presented algorithm is as follows:

Step 0: Let $k = 1$.

Step 1-1: Optimal solution identified as $(u^*, v^*, u_0^*) = (16, 193, 844.05)$ and $P_2^* = P_{16}^* = 1$. Let $G_1 = \{(5.99, 19.49), (6.17, 21.66)\}$, $F_1 : y - 12.05x + 52.68 = 0$ and $k = 2$.

Step 1-2: Optimal solution identified as $(u^*, v^*, u_0^*) = (1, 4, 3.01)$ and $P_{10}^* = P_{16}^* = 1$. Let $G_2 = \{(7.00, 24.99)$
$, (6.17, 21.66)\}$, $F_2 : y - 4.01x + 3.08 = 0$ and $k = 3$.

Step 1-3: Model P_k is infeasible, so the algorithm is terminated and F_k,$k = 1, 2$, are the equations of efficient frontier.

Now we obtain the equations of efficient facet for DMUs with the most input and least output.

Step 0: Let $k = 1$.

Step 1-1: Optimal solution identified as $(u^*, v^*, u_0^*) = (0.73, 0.26, -3.66)$ and $P_{12}^* = P_{14}^* = 1$. Let $G_1 = \{(13.56, 9.82), (14.68, 10.22)\}$, $F_1 : y - 0.35x - 5.08 = 0$ and $k = 2$.

Step 1-2: Optimal solution identified as $(u^*, v^*, u_0^*) = (0.34, 0.65, 5.45)$ and $P_{12}^* = P_{18}^* = 1$. Let $G_2 = \{(13.56, 9.82), (12.77, 8.33)\}$. $F_2 = y - 1.88x + 15.67 = 0$ and $k = 3$.

Step 1-3: Optimal solution identified as $(u^*, v^*, u_0^*) = (0.99, 0.0037, -10.12)$ and $P_4^* = P_{14}^* = 1$. Let $G_3 = \{(12.37, 10.23), (14.68, 10.22)\}$, $F_3 : y - 0.0037x + 10.16 = 0$ and $k = 4$.

Step 1-4: Model P'_k is infeasible, so the algorithm is terminated and $F_k, k = 1, 2, 3$, are the equations of efficient frontier.

5. Conclusion

The suggested method in this paper, finds out efficient frontiers of PPS by solving 0-1 programming problems for DMUs with interval data. By manipulation of the BCC multiplier model, first we obtained equation of efficient frontiers for DMUs with the least input and most output, then for DMUs with the most input and least output, and the number of solved problems is at most equal to the number of efficient facets of PPS. Although performance of algorithm needs more computational effort, but for large n (number of DMUs), solving a little amount of integer problems may be more preferable than solving n linear problems .This is especially important in situations where the number of efficient facets is very less than n.

By means of equations of hyperplans which construct the efficient frontier (weak and strong efficient hyperplanes) we can identify returns to scale of all DMUs and efficiency score of inefficient DMUs in both input and output oriented only with a simple comparison. Therefore, development of algorithm to identifying weak efficient part of PPS provides an interesting challenge for future research.

References

[1] A. I. Ali, Data envelopment analysis: computational issues, computers, Environment and urbon systems 14 (1990) 157-165.

[2] A. I. Ali, streamlined computation for data envelopment analysis, European Journal of Operational Research 64 (1992) 61-67.

[3] A. I. Ali,computational aspects of DEA, In: A. Charnes, W. W. Cooper, A. Y. Lewin, L. M. Seiford, (Eds.), Data Envelopment Analysis: Theory, methodology, And Application, 1994, pp.63-88.

[4] R. D. Banker, A. Charnes, W. W. Cooper, some model for estimating technical and scale inefficiencies in Data Envelopment Analysis, Management science 30 (1984) 1078-1092.

[5] A. Charnes, W. W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European Journal of operational Research 2(6) (1978) 429-444.

[6] W. W. Cooper, K. S, Park, G. Yu, IDEA and AR-IDEA : models for dealing with imprecise data in DEA, Manage. Sci. 45 (1999) 597-607

[7] F. Hosseinzadeh Lotfi,G. R. Jahanshahloo, F. Rezai balf, H. Zhiani rezai, Ranking of DMUs on Interval Data by DEA, Int. J. contemp. Math. Sciences, Vol. 2, 2007, no. 4, 195-201.

[8] F. Hosseinzadeh lotfi, G. R. Jahanshahloo, R. shahverdi, M. Rostami-Malkhalifeh, Cost Efficiency and Cost Malmquist Productivity Index With interval Data, International Mathematical Forum. 2, 2007, no. 9, 441-453.

[9] F. Hosseinzadeh Lotfi, M. Navabakhsh, A. Tehranian, M. Rostami-malkhalifeh, R. Shahverdi, Ranking Bank Branches With Interval Data, The Application Of DEA, International Mathematical Forum, 2, 2007, no. 9, 429-440.

[10] G. R. Jahanshahloo, F. Hosseinzadeh Lotfi, M. Zohrehbandian, Finding the piecewise linear frontier production function in Data Envelopment Analysis, Applied Mathematics and computation 163(1) (2005) 483-488.

[11] P. Korhonen, Searching the Efficient Frontier in Data Envelopment Analysis, IIASA, IR-97-79, 1997.

Received: November 18, 2006